

# Identification of CO and NO<sub>2</sub> using a thermally resistive microsensor and support vector machine

S. Al-Khalifa, S. Maldonado-Bascón and J.W. Gardner

**Abstract:** Levels of carbon monoxide and nitrogen dioxide in air are currently monitored using two different thick-film resistive gas sensors. The resultant high power consumption of thick-film-based gas sensors is problematic for portable multi-gas monitors. The use of a single low-power thermally-modulated resistive gas sensor to monitor simultaneously both gases is reported. The silicon micromachined substrate not only reduces the DC power consumption to 100 mW at 300°C but also permits AC temperature modulation through a small thermal mass. Uniquely, a support vector machine is employed to classify the wavelet coefficients of the AC resistive signal. This simple method permits the rapid classification of CO/NO<sub>2</sub> gas mixtures with a high level of confidence (94% or better) using just one low-power gas microsensor. Thus demonstrating the potential application of a single low-power thermally-modulated resistive gas sensor in portable multi-gas monitors.

## 1 Introduction

Various signal processing methods for use with thermally-modulated sensors for gas sensing, have been reported in the literature [1–6]. For example in [3] a heating element was driven by a sinusoidal voltage and the resulting modulation in sensor resistance analysed by the fast Fourier transform (FFT) method. The amplitudes of the higher harmonics contain information relating to the analyte type and concentration. More recently, these Fourier coefficients were used to train a back-propagation neural network, which was then employed to predict the concentrations of CO and NO<sub>2</sub> in binary mixtures [6].

In 2001, we reported upon the use of a wavelet transform to predict the presence of CO and NO<sub>2</sub> gases in air (e.g. for an alarm) as well as predictions of the actual gas concentrations [7]. Here, we propose the use of a support vector machine (SVM) method to predict more efficiently the concentrations of CO and NO<sub>2</sub> in air using a single thermally-modulated micromachined metal oxide gas sensor.

A SVM classifies into two different groups (classes) objects characterised by a vector ( $\mathbf{x}$ ) of its features. In this work, the feature vector is formed with the wavelet coefficients obtained from the sensor signal. An eight-order Daubechies filter is used to obtain these wavelet coefficients [8].

We now illustrate the simplest case: linear machines trained on separable data. The training data are  $\{\mathbf{x}_i, y_i\}_{i=1}^{N_t}$  ( $N_t$  is the number of training samples),  $\mathbf{x}_i \in \mathbb{R}^d$ —

( $d$  is number of features in the vector) and  $y_i \in \{-1, 1\}$  ('-1' for one class and '1' for the other). In the separable case, a hyperplane,  $\mathbf{w}$  separates positive and negative samples. Those data  $\mathbf{x}$  that satisfy  $\mathbf{w}\mathbf{x} + b = 0$  lie on the hyperplane where  $\mathbf{w}$  is normal to the hyperplane,  $b/\|\mathbf{w}\|$  is the perpendicular distance from the hyperplane to the origin and  $\|\mathbf{w}\|$  is the Euclidean norm of  $\mathbf{w}$ . Samples from positive and negative training samples satisfy:

$$\left. \begin{array}{l} \mathbf{x}_i \mathbf{w} + b \geq 1 \quad \text{for } y_i = 1 \\ \mathbf{x}_i \mathbf{w} + b \leq -1 \quad \text{for } y_i = -1 \end{array} \right\} y_i (\mathbf{x}_i \mathbf{w} + b) - 1 \geq 0 \quad \forall i \quad (1)$$

The SVM creates the separating hyperplane with the largest margin between the hyperplane and any training data.

$$\mathbf{w} = \sum_{i=1}^{N_t} \alpha_i y_i \mathbf{x}_i \quad (2)$$

where the vectors  $\mathbf{x}_i$  with  $\alpha_i = 0$  are the support vectors and the decision function is defined as:

$$f(\mathbf{x}) = \mathbf{w}\mathbf{x} + b = \sum_{i=1}^{N_t} \alpha_i y_i \mathbf{x}_i' + b \quad (3)$$

where  $\alpha_i$  are the Lagrange multipliers from solving the optimisation. One class is assigned when the decision function is positive and the other when negative. A non-linear case can be performed by the transformation of  $\mathbf{x} \rightarrow \phi(\mathbf{x}) \in H$  where the mappings  $\phi(\cdot)$  are performed by a kernel function  $K(\dots)$  that defines an inner product in  $H$ , in this case (3) is modified as follows:

$$f(\mathbf{x}) = \mathbf{w}\phi(\mathbf{x}) + b = \sum_{i=1}^{N_t} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (4)$$

where only  $N_S$  terms must be added ( $N_S$  is the number of support vectors).

In a non-explicit transformation  $\phi(\cdot)$  the plane  $\mathbf{w}$  is not available and the decision function is expressed in terms of the kernel  $K(\dots)$ . The optimal hyperplane is obtained by

© IEE, 2003

IEE Proceedings online no. 20030004

doi:10.1049/ip-smt:20030004

Paper first received 9th April 2002 and in revised form 16th September 2002

S. Al-Khalifa and J.W. Gardner are with the School of Engineering, University of Warwick, Coventry CV4 7AL, UK

S. Maldonado-Bascón is with the Dpto. de Teoría de la Señal y Comunicaciones, Universidad de Alcalá, Alcalá de Henares, Madrid, Spain

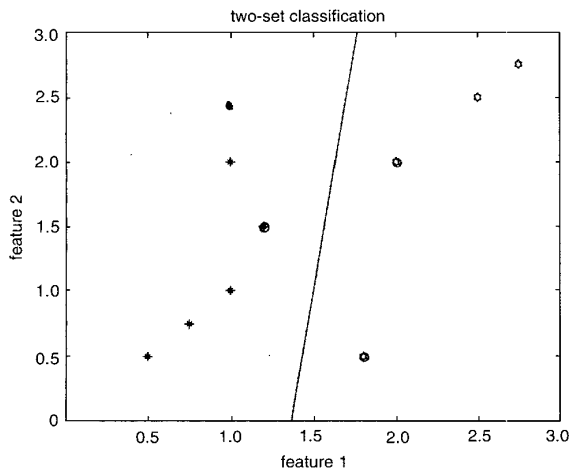
maximising the distance  $L_D$ :

$$L_D = \sum_{i=1}^{N_i} \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_j, x_i) \quad (5)$$

under the constraints of:

$$\sum_{i=1}^{N_i} \alpha_i y_i = 0 \quad \text{and} \quad \alpha_i \geq 0 \quad (6)$$

As can be seen from (5),  $\phi(x)$  is not necessary, so in many cases a kernel  $K(\dots)$  is used without an explicit knowledge of  $\phi(x)$ . A more detailed description of SVMs can be found in [9]. In Fig. 1, two classes have been classified with a linear SVM. The two classes are marked with either a '\*' or a '☆' and the support vectors are circled. In this case the hyperplane ( $w, b$ ) is just a straight (solid) line because of the two-dimensional nature of the problem.



**Fig. 1** Example of the hyperplane classification of data using a SVM machine

Our goal here is to predict the gas type (CO, NO<sub>2</sub> or the mixed CO + NO<sub>2</sub>) and concentration from the response of a thermally-modulated resistive gas microsensor and a SVM.

With the SVM being applied to the wavelet decomposition of the periodic resistive signal [7]. We propose that a SVM can be efficiently applied to gas classification problems. Following training, independent test data are used and then assigned first to one of the two gas classes: C<sub>1</sub> and C<sub>2</sub> (1 or -1). In our implementation,  $x_i$  is formed with the wavelet coefficients obtained from the sensor signal and  $y_i$  is either '1' for 'CO' or '-1' for the other gases at the first stage and '1' for NO<sub>2</sub> and '-1' for the binary gas CO + NO<sub>2</sub> at the second stage. So two SVMs are applied to obtain the gas type. The extension of the classification scheme to gas concentrations is described in Section 4.

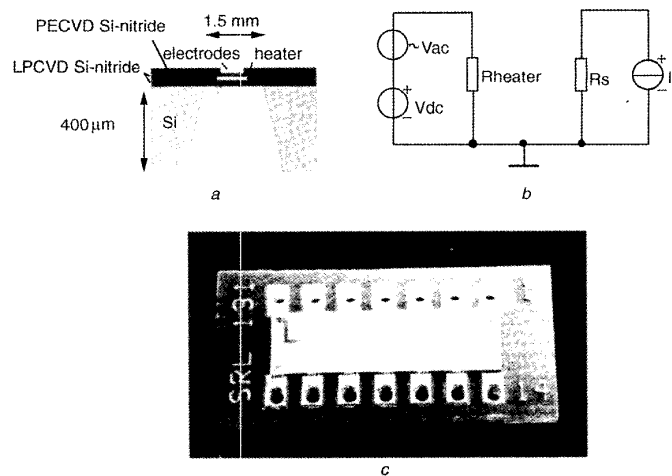
## 2 Gas Testing Station

Fig. 2 shows a schematic cross-section of the silicon micromachined resistive gas sensor [6]. The platinum resistive heating element  $R_{\text{heater}}$  driven by a sinusoidal AC voltage  $V_{\text{ac}}$  with a DC bias  $V_{\text{dc}}$ . The resistance  $R_s$  of a palladium doped tin oxide film lying on top of gold electrodes is measured using a constant current source  $I_s$ . Fig. 2c shows a photograph of an array of three identical devices that were created by depositing the metal oxide film onto the silicon die that was then annealed at 600°C before being bonded onto a 16 pin PCB package.

This gas testing station was controlled by Lab-View software running on a 600 MHz PC with a National Instruments data acquisition card to drive/sense the sensor chip. Full details are presented in [10].

## 3 Feature selection

In many cases, a SVM uses a large number of features for each vector. The most important features must then be selected. In most supervised learning problems, feature selection is important for a variety of reasons: generalisation performance, running time requirements, and constraints and interpretational issues imposed by the problem itself. The feature selection problem addressed here is that of discriminating between finite point sets (i.e. the two classes C<sub>1</sub> and C<sub>2</sub>) in a  $d$ -dimensional feature  $R^d$  by a separating plane that utilises fewer features than  $d$ .



**Fig. 2**  
*a* Cross-section of the silicon micromachined resistive gas sensor,  
*b* Schematic representation of heater drive circuit and sensor measurement circuit  
*c* Photograph of the resistive device used for experimental work

From our experimental procedure, the initial feature vector has 102 features, namely the 102 wavelet coefficients from a four-stage decomposition. In the SVM algorithm, we assume that the mapping function  $\phi(x)$  is available. So, the algorithm can be considered to be modified in such a way that the new training set is  $\{x_i, y_i\} = \{\phi(x_i), y_i\}$  and a hyperplane obtained by its normal  $w$  which is used to classify  $x$ . Classification is performed by the following inequality:

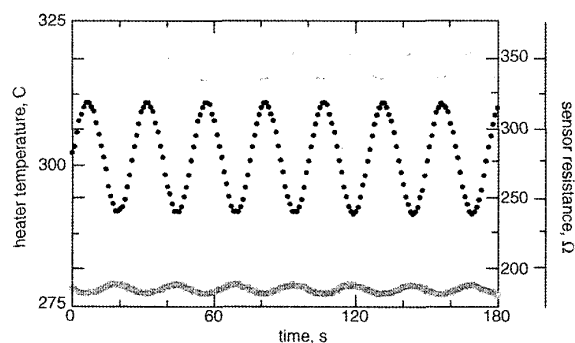
$$x \begin{cases} \text{if } \bar{x}_i, w + b \geq 0 \Rightarrow \text{class 1} \\ \text{if } \bar{x}_i, w + b < 0 \Rightarrow \text{class 2} \end{cases} \quad (7)$$

The angle  $\vartheta_i$  of every feature vector component with the normal of the hyperplane can be obtained. An orthogonal feature component to  $w$  implies an irrelevant feature.

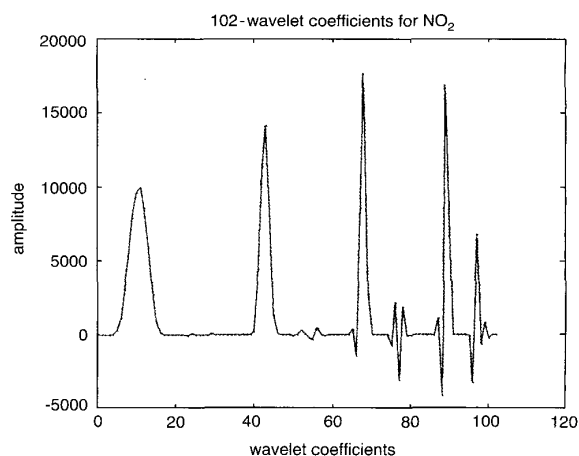
#### 4 Identification of gas type and concentration

##### 4.1 Gas type

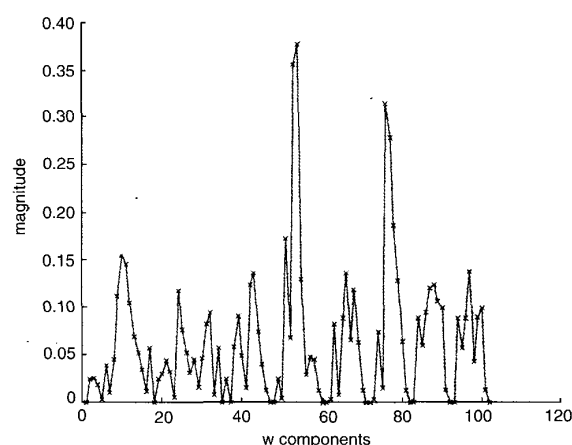
Fig. 3 shows the typical resistive response of the micro-sensor to a modulated heater voltage at two concentrations of CO in air. Here it is only necessary to use just one period of signal (unlike in FFT) to predict the gas type (or concentration). The computation is as follows: one period of the input signal is introduced into a wavelet decomposition block. Four decompositions are performed over the low pass filter output. A single vector is produced concatenating the wavelet coefficients of the four decompositions. The resulting vector for a single period input signal is illustrated in Fig. 4. From the previous decomposition, 102 wavelet coefficients are obtained in total. These 102 wavelet coefficients are used first for gas type classification and later for gas concentration prediction. However, 102 wavelet coefficients present a very high dimensionality and, clearly, many do not contribute significantly to the prediction process. So here, we have used the SVM both for classification and for feature selection. Out of the 102 wavelet coefficients, the coefficients with the most information about the gas are selected. From the SVM, a plane  $w$  that separates two classes is obtained. If the plane is orthogonal to a vector component, then this component is irrelevant. Using this premise, we can eliminate all the components with a scalar product with  $w$  near to zero, i.e. the vector components of low value can be eliminated without reduction of the performance of the classification task. In Fig. 5, the magnitude of the result from training  $w$  components are plotted. These components have been obtained with a reduced training set where one class corresponds to CO and the other class corresponds to



**Fig. 3** Response of device to the AC temperature wave (open squares) in 5 ppm (solid squares) and 80 ppm (circles) of CO in clean air at 25% rh. The temperature wave is only approximately sinusoidal given a sinusoidal drive voltage



**Fig. 4** Graph showing a typical set of wavelet coefficients for the analysis of NO<sub>2</sub> in air



**Fig. 5** Full set of 102 wavelet components used in vector  $w$

CO+NO<sub>2</sub> and NO<sub>2</sub>. The three classes CO, NO<sub>2</sub> and the mixture are separated in two stages. First CO is one class whilst the NO<sub>2</sub> and the mixture are the other class. For this second class another classification is performed to identify between pure NO<sub>2</sub> and the mixture. As can be seen in Fig. 5, there is a high correlation between adjacent coefficients. So, we have selected the maxima from Fig. 5 in order to reduce the correlation between the selected coefficients to classify the gas. Next, the ten largest coefficients are taken from Fig. 5. The result is a reduction in the number of features, and these are selected as the features for the classification. Fig. 6 represents the magnitude of the  $w$  components for the new training process. Using these selected components of the wavelet coefficient vector, tests have been performed on over 182 sets. The gas type classification results are presented in Table 1, and show that the gas type is predicted with an accuracy of over 94%.

##### 4.2 Gas concentration

The SVM can also be used for regression and so predict the gas concentration, a detailed description of SVM regression can be found in [11]. In this case the output  $y_i$  is not binary; instead  $y_i$  is the value of the regressed function for the given input  $x$ . A Gaussian kernel has been used for this part, and is given by

$$K(x, x') = e^{-\sigma|x-x'|^2} \quad (8)$$

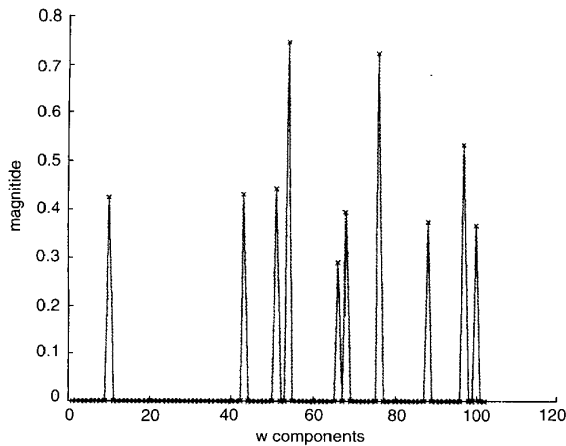


Fig. 6 Reduced set of 10 wavelet components used in vector  $w$

Table 1: Actual classification percentages for the pure CO, pure NO<sub>2</sub> and a mixture of both in air

CO, %	NO <sub>2</sub> , %	CO+NO <sub>2</sub> , %
100	100	94

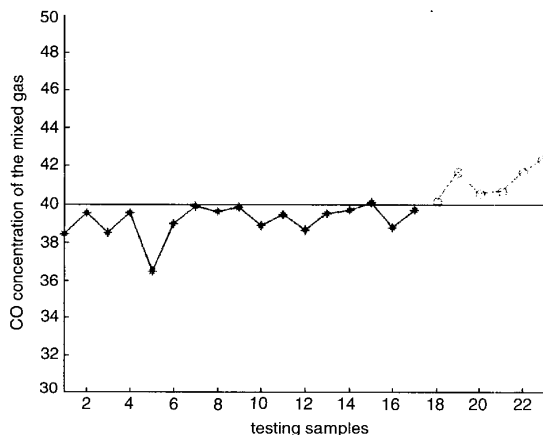


Fig. 7 Concentration of CO in a binary gas mixture based on the training data (labelled \*) and predicted on test data (labelled 0). The nominal gas concentration is 40 ppm, although there will be some small experimental error on this value

where  $\sigma$  is selected by observation of several runs and  $x$ ,  $x'$  are the input and support vectors respectively. Once the gas type is known a different SVM is applied to CO, NO<sub>2</sub> and for the mixed gas. The number of support vectors obtained with the training set of the mixture are 42 for the CO and 33 for the NO<sub>2</sub>. An example of the results is shown in Fig. 7 and this corresponds to a CO concentration of 40 ppm in air. The median results from four different data sets, where six samples have been taken are presented in Table 2, and show the gas concentration can also be predicted accurately with errors of less than 5%.

Table 2: Estimation (prediction) of gas concentration and the associated error

Gas	Real concentration, ppm/ppm	Estimated Concentration, ppm/ppm	Error, %
CO/NO <sub>2</sub>	20/10	20.6/9.46	3.0/5.7
CO/NO <sub>2</sub>	40/20	41.2/19.22	3.0/4.1
CO/NO <sub>2</sub>	80/40	79.0/39.40	1.3/1.5
CO/NO <sub>2</sub>	130/60	123.3/57.22	5.4/4.9

## 5 Conclusions

It has been shown that a single thermally-modulated resistive gas sensor can be employed to predict the presence and levels of CO and NO<sub>2</sub> in air. A simple SVM method can be used to classify the wavelet coefficients and predict unknown gas types and concentrations with a high level of confidence (94% or better). The SVM code can be implemented very efficiently using, for example, a low-cost 8-bit microcontroller and 1 kByte of RAM. Moreover, the ability to detect two pollutant gases in air with one low-power sensor reduces the overall power consumption to a level suitable for commercial battery-powered, multi-gas monitors.

## 6 Acknowledgments

The authors wish to thank Dr. Nicolae Bărsan (Tübingen University) for the deposition of the metal oxide films and Professor Nice de Rooij (Institute of Microtechnology, Neuchatel) for the processing of the silicon devices.

## 7 References

- CAVICCHI, R.E., SUEHLE, J.S., KREIDER, K.G., GAITAN, M., and CHAPARALA, : 'Optimised temperature pulse sequences for the enhancement of chemically-specific response patterns from micro-hotplate gas sensors'. Conf. Proc. Transducers '95 and Eurosensors IX, Stockholm, Sweden, 1995, pp. 823-826
- GÖPEL, W., and SCHIERBAUM, K.D.: 'SnO<sub>2</sub> Sensors: current status and future prospects', *Sens. Actuators B, Chem.*, 1995, **26-27**, pp. 1-12
- NAKATA, S., AKABE, S., NAKASUJI, M., and YOSHIIKAWA, K.: 'Gas sensing based on non-linear response: Discrimination between hydrocarbon and quantification of individual components in a gas mixture', *Anal. Chem.*, 1996, **68**, pp. 2067-2072
- IHOKURA, K., and WATSON, J.: 'Stannic Oxide Gas Sensors, Principles and Applications' (CRC Press, Boca Raton, FL, 1994)
- AL-KHALIFA, S., GARDNER, J.W., and CRAINE, J.F.: 'Characterisation of a thermal wave microsensor for intelligent analysis of atmospheric gases'. Presented at Sensors and their applications VIII, Glasgow, IoP, 1997
- HEILIG, A., BÄRSAN, N., WEIMAR, U., SCHWEIZER-BERBERICH, M., GARDNER, J.W., and GÖPEL, W.: 'Gas identification by modulating temperatures of SnO<sub>2</sub>-based thick film sensors', *Sens. Actuators B, Chem.*, 1997, **43**, pp. 45-51
- AL-KHALIFA, S., GARDNER, J.W., and MALDONADO-BASCÓN, S.: 'Rapid multicomponent analysis using a thermally-modulated resistive gas microsensor and a discrete wavelet transform'. Sensors and their Applications, London, XI/ISMCR, IoP, 2001
- MALLAT, S.: 'A Wavelet tour of signal processing' (Academic Press, 1998)
- BURGES, C.J.C.: 'A tutorial on support vector machines for pattern recognition', *Data Min. Knowl. Discov.*, 1998, **2**, (2), pp. 121-167
- AL-KHALIFA, S.: PhD Thesis, University of Warwick, Coventry CV4 7AL, U.K. 2000
- SMOLA, A.J., and SCHÖLKOPF: 'A tutorial on support vector regression'. Neurocolt2 Technical Report Series, October 1998., (Available at <http://www.neurocolt.com>)